

Blackwell Approachability and Low-Regret Learning are Equivalent

Jacob Abernethy
Computer Science Division
University of California, Berkeley
jake@cs.berkeley.edu

Peter L. Bartlett
Computer Science Division and Department of Statistics
University of California, Berkeley
bartlett@cs.berkeley.edu

Elad Hazan
Faculty of Industrial Engineering & Management
Technion - Israel Institute of Technology
ehazan@ie.technion.ac.il

Abstract

We consider the celebrated Blackwell Approachability Theorem for two-player games with vector payoffs. We show that Blackwell's result is equivalent, via efficient reductions, to the existence of “no-regret” algorithms for Online Linear Optimization. Indeed, we show that any algorithm for one such problem can be efficiently converted into an algorithm for the other. We provide a useful application of this reduction: the first *efficient* algorithm for calibrated forecasting.

1 Introduction

A typical assumption in game theory, and indeed in most of economics, is that an agent’s goal is to optimize a scalar-valued payoff function—a person’s wealth, for example. Such scalar-valued utility functions are the basis for much work in learning and Statistics too, where one hopes to maximize prediction accuracy or minimize expected loss. Towards this end, a natural goal is to prove a guarantee on some algorithm’s minimum expected payoff (or maximum reward).

In 1956, David Blackwell posed an intriguing question: what guarantee can we hope to achieve when playing a two-player game with a *vector-valued payoff*, particularly when the opponent is potentially an adversary? For the case of scalar payoffs, as in a two-player zero-sum game, we already have a concise guarantee by way of Von Neumann’s minimax theorem: either player has a fixed oblivious strategy that is effectively the “best possible”, in that this player could do no better even with knowledge of the opponent’s randomized strategy in advance. This result is equivalent to strong duality for linear programming.

When our payoffs are non-scalar quantities, it does not make sense to ask “can we earn at least x ?”. Instead, we would like to ask “can we guarantee that our vector payoff lies in some convex set S ”? In this case, the story is more difficult, and Blackwell observed that an oblivious strategy does not suffice—in short, we do not achieve “duality” for vector-payoff games. What Blackwell was able to prove is that this negative result applies only for *one-shot games*. In his celebrated Approachability Theorem [3], one can achieve a duality statement *in the limit* when the game is played repeatedly, where the player may learn from his opponent’s prior actions. Blackwell actually constructed an algorithm (that is, an adaptive strategy) with the guarantee that the average payoff vector “approaches” S , hence the name of the theorem.

Blackwell Approachability has the flavor of learning in repeated games, a topic which has received much interest. In particular, there are a wealth of recent results on so-called *no-regret learning algorithms* for making repeated decisions given an arbitrary (and potentially adversarial) sequence of cost functions. The first no-regret algorithm for a “discrete action” setting was given in a seminal paper by James Hannan in 1956 [10]. That same year, David Blackwell pointed out [2] that his Approachability result leads, as a special case, to an algorithm with essentially the same low-regret guarantee proven by Hannan.

Blackwell thus found an intriguing connection between repeated vector-payoff games and low-regret learning, a connection that we shall explore in greater detail in the present work. Indeed, we will show that the relationship goes much deeper than Blackwell had originally supposed. We prove that, in fact, Blackwell’s Approachability Theorem is equivalent, in a very strong sense, to no-regret learning, for the particular setting of so-called “Online Linear Optimization”. Precisely, we show that any no-regret algorithm can be converted into an algorithm for Approachability and vice versa. This algorithmic equivalence is achieved via the use of *conic duality*: if our goal is low-regret learning in a cone K , we can convert this into a problem of approachability of the dual cone K^0 , and vice versa.

This equivalence provides a range of benefits and one such is “calibrated forecasting”. The goal of a calibrated forecaster is to ensure that sequential probability predictions of repeated events are “unbiased” in the following sense: when the weatherman says “30% chance of rain”, it should actually rain roughly three times out of ten. The problem of calibrated forecasting was reduced to Blackwell’s Approachability Theorem by Foster [7], and a handful of other calibration techniques have been proposed, yet none have provided any efficiency guarantees on the strategy. Using a similar reduction from calibration to approachability, and by carefully constructing the reduction from approachability to online linear optimization, we achieve the first efficient calibration algorithm.

Related work There is by now vast literature on all three main topics of this paper: approachability, online learning and calibration, see [4] for an excellent exposition. The relation between the three areas is not as well-understood.

Blackwell himself noted that approachability implies no regret algorithms in the discrete setting. However, as we show hereby, the full power of approachability extends to a much more general framework of online linear optimization, which has only recently been explored (see [12] for a survey) and shown to give the first efficient algorithms for a host of problems (e.g. [1, 6]). Perhaps more significant, we also prove the reverse direction - online linear optimization exactly captures the power of approachability. Previously, it was considered by many to be strictly stronger than regret minimization.

Calibration is a fundamental notion in prediction theory and has found numerous applications in economics and learning. Dawid [5] was the first to define calibration, with numerous algorithms later given by Foster and Vohra [8], Fudenberg and Levine [9], Hart and Mas-Colell [11] and more. Foster has given a calibration algorithm based on approachability [7]. There are numerous definitions of calibration in the literature, mostly asymptotic. In this paper we give precise finite-time rates of calibration and show them to be optimal. Furthermore, we give the first *efficient* algorithm for calibration: attaining ε -calibration (formally defined later) required a running time of $\text{poly}(\frac{1}{\varepsilon})$ for all previous algorithms, whereas our algorithm runs in time proportional to $\log \frac{1}{\varepsilon}$.

2 Preliminaries

2.1 Blackwell Approachability

A *vector-valued game* is defined by a pair of convex compact sets $\mathcal{X} \subset \mathbb{R}^n, \mathcal{Y} \subset \mathbb{R}^m$ and a biaffine mapping $\ell : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}^d$; that is, for any $\alpha \in [0, 1]$ and any $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{X}, \mathbf{y}_1, \mathbf{y}_2 \in \mathcal{Y}$, we have $\ell(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2, \mathbf{y}) = \alpha \ell(\mathbf{x}_1, \mathbf{y}) + (1 - \alpha) \ell(\mathbf{x}_2, \mathbf{y})$, and $\ell(\mathbf{x}, \alpha \mathbf{y}_1 + (1 - \alpha) \mathbf{y}_2) = \alpha \ell(\mathbf{x}, \mathbf{y}_1) + (1 - \alpha) \ell(\mathbf{x}, \mathbf{y}_2)$. We consider $\ell(\mathbf{x}, \mathbf{y})$ to be the “payoff vector” when Player 1 plays strategy \mathbf{x} and Player 2 plays strategy \mathbf{y} . We consider this game from the perspective of Player 1, whom we will often refer to as “the player”, while we refer to Player 2 as “the adversary”.

In a scalar-valued game, the natural question to ask is “how much can a player expect to gain/lose (in expectation) against a worst-case adversary?” With d -dimensional payoffs, of course, we don’t have a notion of ‘more’ or ‘less’, and hence this question does not make sense. As Blackwell pointed out [3], the natural question to consider is “can we guarantee that the payoff vector lies in a given (convex) set S ?” Notice that this formulation dovetails nicely with the original goal in scalar-valued game, in the following way. Take any *halfspace* $H \subset \mathbb{R}^d$, where H is parameterized by a vector \mathbf{v} and a constant c , namely $H = \{\mathbf{z} \in \mathbb{R}^d : \mathbf{z} \cdot \mathbf{v} \geq c\}$. Then the question “can we guarantee that $\ell(\cdot, \cdot)$ lies in H ?” is equivalent to “can the player expect to gain at least c in the scalar-valued game defined by $\ell'(\mathbf{x}, \mathbf{y}) := \ell(\mathbf{x}, \mathbf{y}) \cdot \mathbf{v}$?”

Given that we would like to receive payoff vectors that lie within S , let us define three separate notions of achievement towards this goal.

Definition 1. Let S be a convex set of \mathbb{R}^d .

- We say that a set S is *satisfiable* if there exists a strategy $\mathbf{x} \in \mathcal{X}$ such that for any $\mathbf{y} \in \mathcal{Y}$, $\ell(\mathbf{x}, \mathbf{y}) \in S$.
- We say that a set S is *halfspace-satisfiable* if, for any halfspace $H \supseteq S$, H is satisfiable.
- We say that a set S is *response-satisfiable* if, for any $\mathbf{y} \in \mathcal{Y}$, there exists a $\mathbf{x}_{\mathbf{y}} \in \mathcal{X}$ such that $\ell(\mathbf{x}_{\mathbf{y}}, \mathbf{y}) \in S$.

Among these three conditions the first, satisfiability, is the strongest. Indeed, it says that the player has an *oblivious* strategy which always provides the desired guarantee, namely that the payoff is in S . The second condition, response-satisfiability, is much weaker and says we can achieve the same guarantee provided we observe the opponent’s strategy in advance. We will also make use of the final condition,

halfspace-satisfiability, which is also a weak condition, although we shall show it is equivalent to response-satisfiability.

Of course, a scalar-valued game is a particular case of a vector-valued game. What is interesting is that, for this special case, the condition of satisfiability is in fact no stronger than response-satisfiability for the case when S has the form $[c, \infty)$. Indeed, this fact can be view as the celebrated Minimax Theorem.

Theorem 1 (Von Neumann’s Minimax Theorem [14]). *For \mathcal{X} and \mathcal{Y} the n -dimensional and m -dimensional probability simplexes, and with scalar-valued $\ell(\cdot, \cdot)$, the set $S = [c, \infty)$ is satisfiable if and only if it is response-satisfiable.*

We will also make use of a more general version of the Minimax Theorem, due to Maurice Sion.

Theorem 2 (Sion, 1958 [15]). *Given convex compact sets $\mathcal{X} \subset \mathbb{R}^n$, $\mathcal{Y} \subset \mathbb{R}^m$, and a function $f : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ convex and concave in its first and second arguments respectively, we have*

$$\inf_{\mathbf{x} \in \mathcal{X}} \sup_{\mathbf{y} \in \mathcal{Y}} f(\mathbf{x}, \mathbf{y}) = \sup_{\mathbf{y} \in \mathcal{Y}} \inf_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}, \mathbf{y})$$

One might hope that the analog of Theorem 1 for vector-valued games would also hold true. Unfortunately, this is not the case. Consider the following easy example: $\mathcal{X} = \mathcal{Y} := [0, 1]$, the payoff is simply $\ell(x, y) := (x, y)$ for $x, y \in [0, 1]$, and the set in question is $S := \{(z, z) \mid \forall z \in [0, 1]\}$. Response-satisfiability is easy to establish, simply use the response strategy $x_y = y$. But satisfiability can not be achieved: there is certainly no generic x for which $(x, y) \in S$ for all y .

At first glance, it seems unfortunate that we can not achieve a similar notion of duality for games with vector-valued payoffs. What Blackwell showed, however, is that the story is not quite so bad: we *can* obtain a version of Theorem 1 for a weaker notion of satisfiability. In particular, Blackwell proved that, so long as we can play this game repeatedly, then there exists an adaptive algorithm for playing this game that guarantees satisfiability for the *average* payoff vector in the limit. Blackwell coined the term *approachability*.

Definition 2. *Consider a vector-valued game $\ell(\cdot, \cdot)$ and a convex set S . Imagine we have some “learning” algorithm \mathcal{A} which, given a sequence $\mathbf{y}_1, \mathbf{y}_2, \dots \in \mathcal{Y}$, produces a sequence $\mathbf{x}_1, \mathbf{x}_2, \dots$ via the rule $\mathbf{x}_t \leftarrow \mathcal{A}(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{t-1})$. For any T define¹ the distance of \mathcal{A} to be*

$$D_T(\mathcal{A}; S, \mathbf{y}_1, \dots, \mathbf{y}_T) \equiv \text{dist} \left(\frac{1}{T} \sum_{t=1}^T \ell(\mathbf{x}_t, \mathbf{y}_t), S \right),$$

where here $\text{dist}(\cdot, \cdot)$ mean the usual notion ℓ_2 -distance between a point and a set. For a given vector-valued game and a convex set S , we say that S is approachable if there exists a learning algorithm \mathcal{A} such that,

$$\limsup_{T \rightarrow \infty} D_T(\mathcal{A}; S, \mathbf{y}_1, \dots, \mathbf{y}_T) = 0 \quad \text{for any sequence } \mathbf{y}_1, \mathbf{y}_2, \dots \in \mathcal{Y}$$

Approachability is a curious property: it allows the player to repeat the game and learn from his opponent, and only requires that the average payoff satisfy the desired guarantee in the long run. Blackwell showed that response-satisfiability, which does not imply satisfiability, does imply approachability.

Theorem 3 (Blackwell’s Approachability Theorem [3]). *Any closed convex set S is approachable if and only if it is response-satisfiable.*

This version of the theorem, which appears in Evan-Dar et al. [6], is not the one usually attributed to Blackwell, although this is essentially one of his corollaries. His main theorem states that halfspace-satisfiability, rather than response-satisfiability, implies satisfiability. However, these two weaker satisfiability conditions are equivalent:

¹We may simply write $D_T(\mathcal{A})$ when S and $\mathbf{y}_1, \dots, \mathbf{y}_T$ are clear from context.

Lemma 1. *Given a biaffine function $\ell(\cdot, \cdot)$ and any closed convex set S , S is response-satisfiable if and only if it is halfspace-satisfiable.*

Proof. We will show each direction separately.

\Rightarrow Assume that S is response-satisfiable. Hence, for any \mathbf{y} there is an \mathbf{x}_y such that $\ell(\mathbf{x}_y, \mathbf{y}) \in S$. Now take any halfspace $H \supset S$ parameterized by $\boldsymbol{\theta}, c$, that is $H = \{\mathbf{z} : \langle \boldsymbol{\theta}, \mathbf{z} \rangle \leq c\}$. Then let us define a scalar-valued game with payoff function

$$f(\mathbf{x}, \mathbf{y}) = \langle \boldsymbol{\theta}, \ell(\mathbf{x}, \mathbf{y}) \rangle.$$

Notice that $H \supset S$ implies that $\boldsymbol{\theta} \cdot \mathbf{z} \leq c$ for all $\mathbf{z} \in S$. Combining this with the definition of response-satisfiability, we see that

$$\sup_{\mathbf{y} \in \mathcal{Y}} \inf_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}, \mathbf{y}) \leq \sup_{\mathbf{y} \in \mathcal{Y}} f(\mathbf{x}_y, \mathbf{y}) \leq c.$$

By Sion's Theorem (Theorem 2), it follows that $\inf_{\mathbf{x} \in \mathcal{X}} \sup_{\mathbf{y} \in \mathcal{Y}} f(\mathbf{x}, \mathbf{y}) \leq c$. By compactness of \mathcal{X} , we can choose the minimizer \mathbf{x}^* of this optimization. Notice that, for any $\mathbf{y} \in \mathcal{Y}$, we have that $f(\mathbf{x}^*, \mathbf{y}) \leq c$ by construction, and hence $\ell(\mathbf{x}^*, \mathbf{y}) \in H$. H is thus satisfiable, as desired.

\Leftarrow Assume that S is not response-satisfiable. Hence, there must exist some \mathbf{y}_0 such that $\ell(\mathbf{x}, \mathbf{y}_0) \notin S$ for every \mathbf{x} . Consider the set $U := \{\ell(\mathbf{x}, \mathbf{y}_0) \text{ for all } \mathbf{x} \in X\}$ and notice that U is convex since \mathcal{X} is convex and $\ell(\cdot, \mathbf{y}_0)$ is affine. Furthermore, because S is convex and $S \cap U = \emptyset$ by assumption, there must exist some halfspace H dividing the two, that is $S \subset H$ and $H \cap U = \emptyset$. By construction, we see that for any \mathbf{x} , $\ell(\mathbf{x}, \mathbf{y}_0) \notin H$ and hence H is not satisfiable. It follows immediately that S is not halfspace-satisfiable. □

For the sake of simplicity, and for natural connection to the Minimax Theorem, we prefer Theorem 3. However, for certain results, it will be preferable to appeal to the halfspace-satisfiability condition instead.

2.2 Online Linear Optimization

In the setting Online Linear Optimization, the “learner” makes decisions from a bounded convex decision set \mathcal{K} in some Hilbert space. On each of a sequence of rounds, the decision maker chooses a point $\mathbf{x}_t \in \mathcal{K}$, and is then given a linear cost function $\mathbf{f}_t \in \mathcal{F}$, where \mathcal{F} is some bounded set of cost functions, and cost $\langle \mathbf{f}_t, \mathbf{x}_t \rangle$ is paid. The standard measure of performance in this setting, called regret, is defined as follows.

Definition 3. *The regret of learning algorithm L is*

$$\text{Regret}_T(L) = \max_{\mathbf{f}_1, \dots, \mathbf{f}_T \in \mathcal{F}} \left[\sum_{t=1}^T \langle \mathbf{f}_t, \mathbf{x}_t \rangle - \min_{\mathbf{x} \in K} \sum_{t=1}^T \langle \mathbf{f}_t, \mathbf{x} \rangle \right]$$

We say that L is a no-regret learning algorithm when it holds that $\text{Regret}_T(L) = O(\sqrt{T}) = o(T)$.

We state a well-known result:

Theorem 4. *For any bounded decision set $\mathcal{K} \subset \mathcal{H}$, there exists a no-regret algorithm on \mathcal{K} .*

Later in this paper, we shall use the Gradient Descent algorithm of Zinkevich [16]. Ultimately, our goal will be to show that this theorem is equivalent to Theorem 3.

2.3 Convex cones in Hilbert space

Definition 4. A set $X \subset \mathbb{R}^d$ is a cone if it is closed under addition and multiplication by nonnegative scalars. Given any set $K \subset \mathcal{H}$, define $\text{cone}(K) := \{\alpha \mathbf{x} : \alpha \in \mathbb{R}_+, \mathbf{x} \in K\}$, which is a cone in \mathcal{H} . Also, given any set in Hilbert space $C \subset \mathcal{H}$, we can define the polar cone of C as

$$C^0 := \{\boldsymbol{\theta} \in \mathcal{H} : \langle \boldsymbol{\theta}, \mathbf{x} \rangle \leq 0 \text{ for all } \mathbf{x} \in C\}$$

We state a few simple facts on convex sets:

Lemma 2. If C is a convex cone then (1) $(C^0)^0 = C$ and (2) supporting hyperplanes in C^0 correspond to points $\mathbf{x} \in C$, and vice versa. That is, given any supporting hyperplane H of C^0 , H can be written exactly as $\{\boldsymbol{\theta} \in \mathbb{R}^d : \langle \boldsymbol{\theta}, \mathbf{x} \rangle = 0\}$ for some vector $\mathbf{x} \in C$ that is unique up to scaling.

The distance to a cone can conveniently be measure via a “dual formulation,” as we now show.

Lemma 3. For every convex cone C in Hilbert space

$$\text{dist}(\mathbf{x}, C) = \max_{\boldsymbol{\theta} \in C^0, \|\boldsymbol{\theta}\| \leq 1} \langle \boldsymbol{\theta}, \mathbf{x} \rangle \quad (1)$$

We need to measure distance to \mathcal{K} after we make it into a cone via “lifting”.

Lemma 4. Consider a convex set $\mathcal{K} \subseteq \mathcal{H}$ in Hilbert space and $\mathbf{x} \notin \mathcal{K}$. Let $\|\mathcal{K}\| := \max_{\mathbf{y} \in \mathcal{K}} \|\mathbf{y}\|$. Define C to be the cone generated by the lifting of \mathcal{K} , that is $C = \text{cone}(\{1\} \oplus \mathcal{K})$. Then

$$\text{dist}(1 \oplus \mathbf{x}, C) \leq \text{dist}(\mathbf{x}, \mathcal{K}) \leq (1 + \|\mathcal{K}\|) \cdot \text{dist}(1 \oplus \mathbf{x}, C) \quad (2)$$

3 Duality of Approachability and Low-Regret Learning

We recall the notion of an approachability algorithm \mathcal{A} from Section 2.1. Formally, we imagine \mathcal{A} as a function that observes a sequence of opponent plays $\mathbf{y}_1, \dots, \mathbf{y}_{t-1} \in \mathcal{Y}$ and chooses $\mathbf{x}_t \leftarrow \mathcal{A}(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{t-1})$ from \mathcal{X} , with the goal that $\frac{1}{T} \sum_{t=1}^T \ell(\mathbf{x}_t, \mathbf{y}_t)$ approaches a convex set S . The convex decision sets \mathcal{X}, \mathcal{Y} , the payoff function $\ell(\cdot, \cdot)$, and the set S are all known in advance to \mathcal{A} , and hence we may also write $\mathcal{A}_{\ell, S}$ for the algorithm tuned for these particular choices.

Equivalently, we consider a no-regret algorithm \mathbb{L} as a function that observes a sequence of linear cost functions $\mathbf{f}_1, \dots, \mathbf{f}_{t-1}$ and returns a point $\mathbf{x}_t \leftarrow \mathbb{L}(\mathbf{f}_1, \dots, \mathbf{f}_{t-1})$ from the decision set \mathcal{K} . The goal here is to achieve the regret $\sum_{t=1}^T \langle \mathbf{f}_t, \mathbf{x}_t \rangle - \min_{\mathbf{x} \in \mathcal{K}} \sum_{t=1}^T \langle \mathbf{f}_t, \mathbf{x} \rangle$ that is sublinear in T . The bounded convex set \mathcal{K} is known to the algorithm in advance, and hence we may write $\mathcal{A}_{\mathcal{K}}$ for the algorithm tuned for this particular set \mathcal{K} .

We now prove two claims, showing the equivalence of Blackwell and Online Linear Optimization. Precisely what we will show is the following. Assume we are given (A) an instance of an Online Linear Optimization problem and (B) an algorithm that achieves the goals of Blackwell’s Approachability Theorem. Then we shall show that we can convert the algorithm for (B) to achieve a no-regret algorithm for (A).

Lemma 5. For S defined in Algorithm 1, there exists an approachability algorithm \mathcal{A} for S ; that is, $D_T(\mathcal{A}; S) \rightarrow 0$ as $T \rightarrow \infty$.

Proposition 1. The reduction defined in Algorithm 1, for any input algorithm \mathcal{A} , produces an OLO algorithm \mathbb{L} such that $\frac{\text{Regret}(\mathbb{L})}{T} \leq (1 + \|\mathcal{K}\|) D_T(\mathcal{A})$.

Algorithm 1 Reduction of Approachability Alg. \mathcal{A} to Online Linear Optimization Alg. \mathbb{L}

Input: Convex decision set $\mathcal{K} \subset \mathbb{R}^d$
Input: Sequence of cost functions $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_T \in \mathbb{R}^d$
Input: Approachability algorithm \mathcal{A}
Set: Two-player vector-payoff game $\ell : \mathcal{K} \times \mathbb{R}^d \rightarrow \mathbb{R}^{d+1}$ as $\ell(\mathbf{x}, \mathbf{f}) = \langle \mathbf{f}, \mathbf{x} \rangle \oplus -\mathbf{f}$
Set: Approach set $S := \text{cone}(1 \oplus \mathcal{K})^0$
for $t = 1, \dots, T$ **do**
 Let: $\mathbb{L}(\mathbf{f}_1, \dots, \mathbf{f}_{t-1}) := \mathcal{A}_{\ell, S}(\mathbf{f}_1, \dots, \mathbf{f}_{t-1})$
 Receive: cost function \mathbf{f}_t
end for

Algorithm 2 Conversion of Online Linear Optimization Alg. \mathbb{L} to Approachability Alg. \mathcal{A}

Input: Convex compact decision sets $\mathcal{X} \subset \mathbb{R}^n$ and $\mathcal{Y} \subset \mathbb{R}^m$
Input: Biaffine vector-payoff function $\ell(\cdot, \cdot) : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}^d$
Input: Approaching Set $S \subset \{1\} \times \mathbb{R}^{d-1}$
Input: Online Linear Optimization algorithm \mathbb{L}
Set: $\mathcal{K} = \text{cone}(S)^0 \cap B_1$
for $t = 1, \dots, T$ **do**
 Query: $\theta_t \leftarrow \mathbb{L}_{\mathcal{K}}(\mathbf{f}_1, \dots, \mathbf{f}_{t-1})$, where $\mathbf{f}_s \leftarrow -\ell(\mathbf{x}_s, \mathbf{y}_s)$
 Compute: $\mathbf{x}_t \in \mathcal{X}$ so that $\langle \theta_t, \ell(\mathbf{x}_t, \mathbf{y}) \rangle \leq 0$ for any $\mathbf{y} \in \mathcal{Y}$ *// Halfspace oracle*
 Let: $\mathcal{A}(\mathbf{y}_1, \dots, \mathbf{y}_{t-1}) := \mathbf{x}_t$
 Receive: $\mathbf{y}_t \in \mathcal{Y}$
end for

Now onto the second reduction. The construction in Algorithm 2 attempts the following. Assume we are given (A) an instance of a vector-payoff game and an approach set S and (B) a low-regret OLO algorithm. Then we shall show that we can convert the algorithm for (B) to achieve approachability for (A).

Proposition 2. *The reduction in Algorithm 2 produces an approachability algorithm \mathcal{A} with distance bounded by*

$$D_T(\mathcal{A}) \leq (1 + \|S\|) \frac{\text{Regret}(\mathbb{L})}{T}$$

as long as S is halfspace-satisfiable with respect to $\ell(\cdot, \cdot)$.

4 Efficient Calibration via Approachability and OLO

Imagine a sequence of binary outcomes, say ‘rain’ or ‘shine’ on a given day, and imagine a forecaster, say the weatherman, that wants to predict the probability of this outcome on each day. A natural question to ask is, on the days when the weatherman actually predicts “30% chance of rain”, does it actually rain (roughly) 30% of the time? This exactly the problem of *calibrated forecasting* which we now discuss.

There have been a range of definitions of calibration given throughout the literature, some equivalent and some not, but from a computational viewpoint there are significant differences. We thus give a clean definition of calibration, first introduced by Foster [7], which is convenient to assess computationally.

We let $y_1, y_2, \dots \in \{0, 1\}$ be a sequence of outcomes, and $p_1, p_2, \dots \in [0, 1]$ a sequence of probability predictions by a forecaster. We define for every T and every probability interval $[a, b]$, where $0 \leq a \leq b \leq 1$,

the quantities

$$n_T(p, \varepsilon) := \sum_{t=1}^T \mathbb{I}[p_t \in (p - \varepsilon/2, p + \varepsilon/2)], \quad \rho_T(p, \varepsilon) := \frac{\sum_{t=1}^T y_t \mathbb{I}[p_t \in (p - \varepsilon/2, p + \varepsilon/2)]}{n_T(p, \varepsilon)}.$$

The quantity $\rho_T(p - \varepsilon/2, p + \varepsilon/2)$ should be interpreted as the empirical frequency of $y_t = 1$, up to round T , on only those rounds where the forecaster's prediction was “roughly” equal to p . The goal of calibration, of course, is to have this empirical frequency $\rho_T(p, \varepsilon)$ be close to the estimated frequency p , which leads us to the following definition.

Definition 5. Let the (ℓ_1, ε) -calibration rate for forecaster \mathcal{A} be

$$C_T^\varepsilon(\mathcal{A}) = \sum_{i=0}^{\lfloor \varepsilon^{-1} \rfloor} \frac{n_T(i\varepsilon, \varepsilon)}{T} |i\varepsilon - \rho_T(i\varepsilon, \varepsilon)| - \frac{\varepsilon}{2}$$

We say that a forecaster is (ℓ_1, ε) -calibrated if $C_T^\varepsilon(\mathcal{A}) = o(1)$. This in turn implies $\limsup_{T \rightarrow \infty} C_T^\varepsilon(\mathcal{A}) = 0$.

This definition emphasizes that we can ignore an interval $(p - \varepsilon/2, p + \varepsilon/2)$ in cases when our forecaster “rarely” makes predictions within this interval—more precisely, when we forecast within this interval with a frequency that is sublinear in T . Another important feature of this definition is the constant $\varepsilon/2$ - which is an artifact of the discretization by ε . This is the smallest constant which allows for $\limsup_{T \rightarrow \infty} C_T^\varepsilon(\mathcal{A}) = 0$.

We give an equivalent and alternative characterization of this definition: let the calibration vector at time T denoted c_T be given by: $c_T(i) = \frac{n_T(i\varepsilon, \varepsilon)}{T} |i\varepsilon - \rho_T(i\varepsilon, \varepsilon)|$

Claim 1. The (ℓ_1, ε) -calibration rate is equal to the distance of the calibration vector to the ℓ_1 ball of radius $\varepsilon/2$:

$$C_T^\varepsilon = \text{dist}(c_T, B_1(\varepsilon/2))$$

Proof. Notice that:

$$\text{dist}_1(\mathbf{x}, B_1(\varepsilon/2)) := \min_{\mathbf{y}: \|\mathbf{y}\|_1 \leq \varepsilon/2} \|\mathbf{x} - \mathbf{y}\|_1 = -\varepsilon/2 + \|\mathbf{x}\|_1$$

where the second equality follows by noting that an optimally chosen \mathbf{y} will lie in the same quadrant as \mathbf{x} . \square

A standard reduction in the literature (see e.g. [4]) shows that ε -calibration and full calibration are essentially the same (in the sense that an ε -calibrated algorithm can be converted to a calibrated one). For simplicity we consider only ε -calibration henceforth.

4.1 Existence of Calibrated Forecaster via Blackwell Approachability

A surprising fact is that it is possible to achieve calibration even when the outcome sequence $\{y_t\}$ is chosen by an adversary, although this requires a randomized strategy of the forecaster. Algorithms for calibrated forecasting under adversarial conditions have been given in Foster and Vohra [8], Fudenberg and Levine [9], and Hart and Mas-Colell [11].

Interestingly, the calibration problem was reduced to Blackwell's Approachability Theorem in a short paper by Foster in 1999 [7]. Foster's reduction uses Blackwell's original theorem, proving that a given set is halfspace-satisfiable, in particular by providing a construction for each such halfspace. Here, we provide a reduction to Blackwell Approachability using the response-satisfiability condition, i.e. via Theorem 3,

which is both significantly easier and more intuitive than Blackwell². We also show, using the reduction to Online Linear Optimization from the previous section, how to achieve the most efficient known algorithm for calibration by taking advantage of the Online Gradient Descent algorithm of Zinkevich [16], using the results of Section 3.

We now describe the construction that allows us to reduce calibration to approachability. For any $\varepsilon > 0$ we will show how to construct an (ℓ_1, ε) -calibrated forecaster. Notice that from here, it is straightforward to produce a well-calibrated forecaster [8]. For simplicity, assume $\varepsilon = 1/m$ for some positive integer m . On each round t , a forecaster will now randomly predict a probability $p_t \in \{0/m, 1/m, 2/m, \dots, (m-1)/m, 1\}$, according to the distribution \mathbf{w}_t , that is $\Pr(p_t = i/m) = w_t(i)$. We now define a vector-valued game. Let the player choose $\mathbf{w}_t \in \mathcal{X} := \Delta_{m+1}$, and the adversary choose $y_t \in \mathcal{Y} := [0, 1]$, and the payoff vector will be

$$\ell(\mathbf{w}_t, y_t) := \left\langle \mathbf{w}_t(0) \left(y_t - \frac{0}{m} \right), \mathbf{w}_t(1) \left(y_t - \frac{1}{m} \right), \dots, \mathbf{w}_t(m) (y_t - 1) \right\rangle \quad (3)$$

Lemma 6. *Consider the vector-valued game described above and let S , the ℓ_1 ball of radius $\varepsilon/2$. If we have a strategy for choosing \mathbf{w}_t that guarantees approachability of S , that is $\frac{1}{T} \sum_{t=1}^T \ell(\mathbf{w}_t, y_t) \rightarrow S$, then a randomized forecaster that selects p_t according to \mathbf{w}_t is (ℓ_1, ε) -calibrated with high probability.*

The proof of this lemma is straightforward, and is similar to the construction in Foster [7]. The vector $\frac{1}{T} \sum_{t=1}^T \ell(\mathbf{w}_t, y_t)$ is simply the expectation of (ℓ_1, ε) -calibration vector at T . Since each p_t is drawn independently, by standard concentration arguments we can see that if $\frac{1}{T} \sum_{t=1}^T \ell(\mathbf{w}_t, y_t)$ is close to the ℓ_1 ball of radius $\varepsilon/2$, then the (ℓ_1, ε) -calibration vector is close to the $\varepsilon/2$ ball with high probability.

We can now apply Theorem 3 to prove the existence of a calibrated forecaster.

Theorem 5. *For the vector-valued game defined in (3), the ℓ_1 ball of radius $\varepsilon/2$ is response-satisfiable and, hence, approachable.*

Proof. To show response-satisfiability, we need only show that, for every strategy $y \in [0, 1]$ played by the adversary, there is a strategy $\mathbf{w} \in \Delta_m$ for which $\ell(\mathbf{w}, y) \in S$. This can be achieved by simply setting i so as to minimize $|i\varepsilon - y|$, which can always be made smaller than $\varepsilon/2$. We then choose our distribution $\mathbf{w} \in \Delta_{m+1}$ to be a point mass on i , that is we set $w(i) = 1$ and $w(j) = 0$ for all $j \neq i$. Then $\ell(\mathbf{w}, y)$ is identically 0 everywhere except the i th coordinate, which has the value $y - i/m$. By construction, $y - i/m \in [-1/m, 1/m]$, and we are done. \square

4.2 Efficient Algorithm for Calibration via Online Linear Optimization

We now show how the results in the previous Section lead to the first efficient algorithm for calibrated forecasting. The previous theorem provides a natural existence proof for Calibration, but it does not immediately provide us with a simple and efficient algorithm. We proceed according to the reduction outlined in the previous section to prove:

Theorem 6. *There exists a (ℓ_1, ε) -calibration algorithm that runs in time $O(\log \frac{1}{\varepsilon})$ per iteration and satisfies $C_T^\varepsilon = O\left(\frac{1}{\sqrt{\varepsilon T}}\right)$*

The reduction developed in Proposition 2 has some flexibility, and we shall modify it for the purposes of this problem. The objects we shall need, as well as the required conditions, are as follows:

1. A convex set \mathcal{K}

²A similar existence proof was discovered concurrently by Mannor and Stoltz [13]

2. An efficient learning algorithm \mathcal{A} which, for any sequence $\mathbf{f}_1, \mathbf{f}_2, \dots$, can select a sequence of points $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots \in \mathcal{K}$ with the guarantee that $\sum_{t=1}^T \langle \mathbf{f}_t, \boldsymbol{\theta}_t \rangle - \min_{\boldsymbol{\theta} \in \mathcal{K}} \sum_{t=1}^T \langle \mathbf{f}_t, \boldsymbol{\theta} \rangle = o(T)$. For the reduction, we shall set $\mathbf{f}_t \leftarrow -\ell(\mathbf{w}_t, y_t)$.
3. An efficient oracle that can select a particular $\mathbf{w}_t \in \mathcal{X}$ for each $\boldsymbol{\theta}_t \in \mathcal{K}$ with the guarantee that

$$\text{dist} \left(\frac{1}{T} \sum_{t=1}^T \ell(\mathbf{w}_t, y_t), S \right) \leq \frac{1}{T} \left(\sum_{t=1}^T \langle -\ell(\mathbf{w}_t, y_t), \boldsymbol{\theta}_t \rangle - \min_{\boldsymbol{\theta} \in \mathcal{K}} \sum_{t=1}^T \langle -\ell(\mathbf{w}_t, y_t), \boldsymbol{\theta} \rangle \right) \quad (4)$$

where the function $\text{dist}()$ can be with respect to any norm.

The Setup Let $\mathcal{K} = B_\infty(1) = \{\boldsymbol{\theta} \in \mathbb{R}^d : \|\boldsymbol{\theta}\|_\infty \leq 1\}$ be the unit cube. This is an appropriate choice because we can write the ℓ_1 distance to $B_1(\varepsilon/2)$ (the ℓ_1 ball of radius $\varepsilon/2$) as

$$\text{dist}_1(\mathbf{x}, B_1(\varepsilon/2)) := \min_{\mathbf{y}: \|\mathbf{y}\|_1 \leq \varepsilon/2} \|\mathbf{x} - \mathbf{y}\|_1 = -\varepsilon/2 + \|\mathbf{x}\|_1 = -\varepsilon/2 - \min_{\boldsymbol{\theta}: \|\boldsymbol{\theta}\|_\infty \leq 1} \langle -\mathbf{x}, \boldsymbol{\theta} \rangle, \quad (5)$$

where the second equality follows by noting that an optimally chosen \mathbf{y} will lie in the same quadrant as \mathbf{x} . Furthermore, we shall construct our oracle mapping $\boldsymbol{\theta} \mapsto \mathbf{w}$ with the following guarantee: $\langle \ell(\mathbf{w}, y), \boldsymbol{\theta} \rangle \leq \varepsilon/2$ for any y . Using this guarantee, and if we plug in $\mathbf{x} = \frac{1}{T} \sum_{t=1}^T \ell(\mathbf{w}_t, y_t)$ (5), we arrive at:

$$\begin{aligned} \text{dist}_1 \left(\frac{\sum_{t=1}^T \ell(\mathbf{w}_t, y_t)}{T}, B_1(\varepsilon/2) \right) &= -\varepsilon/2 - \min_{\boldsymbol{\theta}: \|\boldsymbol{\theta}\|_\infty \leq 1} \left\langle \frac{-\sum_{t=1}^T \ell(\mathbf{w}_t, y_t)}{T}, \boldsymbol{\theta} \right\rangle \\ &\leq \frac{1}{T} \left(\sum_{t=1}^T \langle -\ell(\mathbf{w}_t, y_t), \boldsymbol{\theta}_t \rangle - \min_{\boldsymbol{\theta} \in \mathcal{K}} \sum_{t=1}^T \langle -\ell(\mathbf{w}_t, y_t), \boldsymbol{\theta} \rangle \right) \end{aligned}$$

This is precisely the necessary guarantee (4).

Constructing the Oracle We now turn our attention to designing the required oracle in an *efficient* manner. In particular, given any $\boldsymbol{\theta}$ with $\|\boldsymbol{\theta}\|_\infty \leq 1$ we must construct $\mathbf{w} \in \Delta_{m+1}$ so that $\langle \ell(\mathbf{w}, y), \boldsymbol{\theta} \rangle \leq \varepsilon/2$ for any y . The details of this oracle are given in Algorithm 3. It is straightforward why, in the final **else** condition,

Algorithm 3 Constructing \mathbf{w} from $\boldsymbol{\theta}$

Input: $\boldsymbol{\theta}$ such that $\|\boldsymbol{\theta}\|_\infty \leq 1$

if $\theta(0) \leq 0$ **then**

$\mathbf{w} \leftarrow \delta_0$ // That is, choose \mathbf{w} to place all weight on the 0th coordinate

else if $\theta(m) \geq 0$ **then**

$\mathbf{w} \leftarrow \delta_m$ // That is, choose \mathbf{w} to place all weight on the last coordinate

else

Binary search $\boldsymbol{\theta}$ to find coordinate i such that $\theta(i) > 0$ and $\theta(i+1) \leq 0$

$\mathbf{w} \leftarrow \frac{\theta(i)^{-1}}{\theta(i)^{-1} - \theta(i+1)^{-1}} \delta_i + \frac{-\theta(i+1)^{-1}}{\theta(i)^{-1} - \theta(i+1)^{-1}} \delta_{i+1}$

end if

Return \mathbf{w}

there must be such a pair of coordinates $i, i+1$ satisfying the condition. We need not be concerned with the case that $\theta(i+1) = 0$, where we can simply define $\frac{0}{\infty} = 0$ and $\frac{\infty}{\infty} = 1$ leading to $\mathbf{w} \leftarrow \delta_{i+1}$. It is also clear that, with the binary search, this algorithm requires at most $O(\log m) = O(\log 1/\varepsilon)$ computation.

In order to prove that this construction is valid we need to check the condition that, for any $y \in \{0, 1\}$, $\langle \ell(\mathbf{w}, y), \boldsymbol{\theta} \rangle \leq \varepsilon/2$; or more precisely, $\sum_{i=1}^m \boldsymbol{\theta}(i) \mathbf{w}(i) \left(y - \frac{i}{m}\right) \leq \varepsilon/2$. Recalling that $m = 1/\varepsilon$, this is trivially checked for the case when $\boldsymbol{\theta}(1) \leq 0$ or $\boldsymbol{\theta}(m) \geq 0$. Otherwise, we have

$$\begin{aligned} \langle \ell(\mathbf{w}, y), \boldsymbol{\theta} \rangle &= \boldsymbol{\theta}(i) \frac{\boldsymbol{\theta}(i)^{-1}}{\boldsymbol{\theta}(i)^{-1} - \boldsymbol{\theta}(i+1)^{-1}} \left(y - \frac{i}{m}\right) + \boldsymbol{\theta}(i+1) \frac{-\boldsymbol{\theta}(i+1)^{-1}}{\boldsymbol{\theta}(i)^{-1} - \boldsymbol{\theta}(i+1)^{-1}} \left(y - \frac{i+1}{m}\right) \\ &= \frac{1}{\boldsymbol{\theta}(i)^{-1} - \boldsymbol{\theta}(i+1)^{-1}} \frac{1}{m} \leq \frac{\max(|\boldsymbol{\theta}(i)|, |\boldsymbol{\theta}(i+1)|)}{2} \varepsilon \leq \frac{\varepsilon}{2} \end{aligned}$$

The Learning Algorithm The final piece is to construct an efficient learning algorithm which leads to vanishing regret. That is, we need to construct a sequence of $\boldsymbol{\theta}_t$'s in the unit cube (denoted $B_\infty(1)$) so that

$$\sum_{t=1}^T \langle \ell_t, \boldsymbol{\theta}_t \rangle - \min_{\boldsymbol{\theta} \in B_\infty(1)} \sum_{t=1}^T \langle \ell_t, \boldsymbol{\theta} \rangle = o(T),$$

where $\ell_t := \ell(\mathbf{w}_t, y_t)$. There are a range of possible no-regret algorithms available, but we use the one given by Zinkevich known commonly as Online Gradient Descent [16]. The details are given in Algorithm 4. This algorithm can indeed be implemented efficiently, requiring only $O(1)$ computation on each

Algorithm 4 Online Gradient Descent

Input: convex set $\mathcal{K} \subset \mathbb{R}^d$
Initialize: $\boldsymbol{\theta}_1 = \mathbf{0}$
Set Parameter: $\eta = O(T^{-1/2})$
for $t = 1, \dots, T$ **do**
 Receive ℓ_t
 $\boldsymbol{\theta}'_{t+1} \leftarrow \boldsymbol{\theta}_t - \eta \ell_t$ // Gradient Descent Step
 $\boldsymbol{\theta}_{t+1} \leftarrow \text{Project}_2(\boldsymbol{\theta}'_{t+1}, \mathcal{K})$ // L2 Projection Step
end for

round and $O(\min\{m, T\})$ memory. The main advantage is that the vectors ℓ_t are generated via our oracle above, and these vectors are *sparse*, having only at most two nonzero coordinates. Hence, the Gradient Descent Step requires only $O(1)$ computation. In addition, the Projection Step can also be performed in an efficient manner. Since we assume that $\boldsymbol{\theta}_t \in B_\infty(1)$, the updated point $\boldsymbol{\theta}'_{t+1}$ can violate at most two of the L_∞ constraints of the ball $B_\infty(1)$. An ℓ_2 projection onto the cube requires simply rounding the violated coordinates into $[-1, 1]$. The number of non-zero elements in $\boldsymbol{\theta}$ can increase by at most two every iteration, and storing $\boldsymbol{\theta}$ is the only state that online gradient descent needs to store, hence the algorithm can be implemented with $O(\min\{T, m\})$ memory. We thus arrive at an efficient no-regret algorithm for choosing $\boldsymbol{\theta}_t$.

Proof of Theorem 6. Here we have bounded the distance directly by the regret, using equation (4), which tells us that the calibration rate is bounded by the regret of the online learning algorithm. Online Gradient Descent guarantees the regret to be no more than $DG\sqrt{T}$, where D is the ℓ_2 diameter of the set, and G is the ℓ_2 -norm of the largest cost vector. For the ball $B_\infty(1)$, the diameter $D = \sqrt{\frac{1}{\varepsilon}}$, and we can bound the norm of our loss vectors by $G = \sqrt{2}$. Hence:

$$C_T^\varepsilon = \text{dist}(c_T, B_1(\varepsilon/2)) \leq \frac{\text{Regret}_T}{T} \leq \frac{GD}{\sqrt{T}} = O\left(\frac{1}{\sqrt{\varepsilon T}}\right) \quad (6)$$

□

References

- [1] Jacob Abernethy, Elad Hazan, and Alexander Rakhlin. Competing in the dark: An efficient algorithm for bandit linear optimization. In *COLT*, pages 263–274, 2008.
- [2] D. Blackwell. Controlled random walks. In *Proceedings of the International Congress of Mathematicians*, volume 3, pages 336–338, 1954.
- [3] D. Blackwell. An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6(1):18, 1956.
- [4] Nicolo Cesa-Bianchi and Gbor Lugosi. *Prediction, Learning, and Games*. 2006.
- [5] A. Dawid. The well-calibrated Bayesian. *Journal of the American Statistical Association*, 77:605–613, 1982.
- [6] E. Even-Dar, R. Kleinberg, S. Mannor, and Y. Mansour. Online learning for global cost functions. 2009.
- [7] D. P Foster. A proof of calibration via blackwell’s approachability theorem. *Games and Economic Behavior*, 29(1-2):7378, 1999.
- [8] D. P Foster and R. V Vohra. Asymptotic calibration. *Biometrika*, 85(2):379, 1998.
- [9] D. Fudenberg and D. K Levine. An easier way to calibrate* 1. *Games and economic behavior*, 29(1-2):131137, 1999.
- [10] J. Hannan. Approximation to Bayes risk in repeated play. *Contributions to the Theory of Games*, 3:97–139, 1957.
- [11] S. Hart and A. Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):11271150, 2000.
- [12] Elad Hazan. The convex optimization approach to regret minimization. In *To appear in Optimization for Machine Learning*. MIT Press, 2010.
- [13] Shie Mannor and Gilles Stoltz. A Geometric Proof of Calibration. *arXiv*, Dec 2009.
- [14] J. Von Neumann, O. Morgenstern, H. W Kuhn, and A. Rubinstein. *Theory of games and economic behavior*. Princeton university press Princeton, NJ, 1947.
- [15] M. Sion. On general minimax theorems. *Pacific J. Math*, 8(1):171–176, 1958.
- [16] M. Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *MACHINE LEARNING-INTERNATIONAL WORKSHOP THEN CONFERENCE-*, volume 20, page 928, 2003.

A Proofs

Proof of Lemma 3. We need two simple observations. Define $\pi_C(\mathbf{x})$ as the projection of \mathbf{x} onto C . Then clearly, for any \mathbf{x} ,

$$\text{dist}(\mathbf{x}, C) = \|\mathbf{x} - \pi_C(\mathbf{x})\| \quad (7)$$

$$\langle \mathbf{x} - \pi_C(\mathbf{x}), \mathbf{y} \rangle \leq 0 \quad \forall \mathbf{y} \in C \text{ and hence } \mathbf{x} - \pi_C(\mathbf{x}) \in C^0 \quad (8)$$

$$\langle \mathbf{x} - \pi_C(\mathbf{x}), \pi_C(\mathbf{x}) \rangle = 0 \quad (9)$$

Given any $\boldsymbol{\theta} \in C^0$ with $\|\boldsymbol{\theta}\| \leq 1$, since $\pi_C(\mathbf{x}) \in C$ we have that

$$\langle \boldsymbol{\theta}, \mathbf{x} \rangle \leq \langle \boldsymbol{\theta}, \mathbf{x} - \pi_C(\mathbf{x}) \rangle \leq \|\boldsymbol{\theta}\| \|\mathbf{x} - \pi_C(\mathbf{x})\| \leq \|\mathbf{x} - \pi_C(\mathbf{x})\|,$$

which immediately implies that $\max_{\boldsymbol{\theta} \in C^0, \|\boldsymbol{\theta}\| \leq 1} \langle \boldsymbol{\theta}, \mathbf{x} \rangle \leq \text{dist}(\mathbf{x}, C)$. Furthermore, by selecting $\boldsymbol{\theta} = \frac{\mathbf{x} - \pi_C(\mathbf{x})}{\|\mathbf{x} - \pi_C(\mathbf{x})\|}$ which has norm one and, by (7), is in C^0 , we see that

$$\max_{\boldsymbol{\theta} \in C^0, \|\boldsymbol{\theta}\| \leq 1} \langle \boldsymbol{\theta}, \mathbf{x} \rangle \geq \left\langle \frac{\mathbf{x} - \pi_C(\mathbf{x})}{\|\mathbf{x} - \pi_C(\mathbf{x})\|}, \mathbf{x} \right\rangle = \left\langle \frac{\mathbf{x} - \pi_C(\mathbf{x})}{\|\mathbf{x} - \pi_C(\mathbf{x})\|}, \mathbf{x} - \pi_C(\mathbf{x}) \right\rangle = \|\mathbf{x} - \pi_C(\mathbf{x})\|,$$

which implies that $\max_{\boldsymbol{\theta} \in C^0, \|\boldsymbol{\theta}\| \leq 1} \langle \boldsymbol{\theta}, \mathbf{x} \rangle \geq \text{dist}(\mathbf{x}, C)$ and hence we are done. \square

Proof of Lemma 4. Since $\text{dist}(1 \oplus \mathbf{x}, \{1\} \times \mathcal{K}) = \text{dist}(\mathbf{x}, \mathcal{K})$ and $\{1\} \oplus \mathcal{K} \subset C$, the first inequality follows immediately.

For the second inequality, let $\mathbf{y} = 1 \oplus \mathbf{x}$ and let \mathbf{w}, \mathbf{v} be the closest points to \mathbf{y} in C and \mathcal{K} respectively. Consider the plane determined by these three points, as depicted in figure 1. Notice that, by triangle

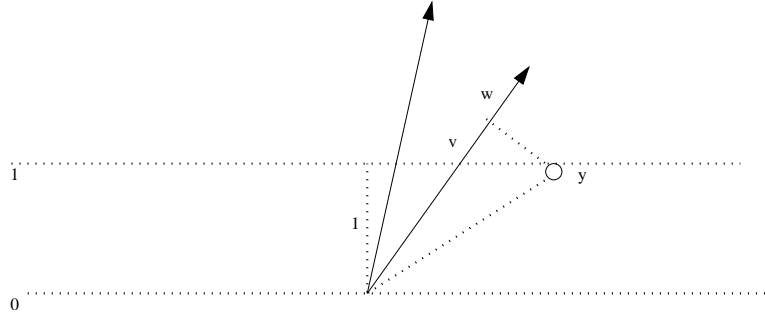


Figure 1: The ratio of distances to \mathcal{K} and the cone is the same as the ratio between $\|\mathbf{v}\|$ and one.

similarity, we have that

$$\|\mathbf{v}\| = \frac{\|\mathbf{v}\|}{\|1 \oplus \mathbf{0}\|} = \frac{\|\mathbf{y} - \mathbf{v}\|}{\|\mathbf{y} - \mathbf{w}\|} = \frac{\text{dist}(\mathbf{y}, \{1\} \oplus \mathcal{K})}{\text{dist}(\mathbf{y}, C)}$$

Of course, $\mathbf{v} \in \mathcal{K}$ and hence $\|\mathbf{v}\| \leq \|\{1\} \oplus \mathcal{K}\| \leq 1 + \|\mathcal{K}\|$. The result follows immediately. \square

Proof of Lemma 5. The existence of an approachability algorithm is established by Blackwell's Approachability Theorem (Theorem 3), as long as we can guaranteed the response-satisfiability condition. Precisely, we must show that, for any \mathbf{f} , there is some $\mathbf{x}_f \in \mathcal{K}$ such that $\ell(\mathbf{x}_f, \mathbf{f}) = \langle \mathbf{f}, \mathbf{x}_f \rangle \oplus -\mathbf{f} \in \text{cone}(1 \oplus \mathcal{K})^0$. Recall that $\theta \in \text{cone}(1 \oplus \mathcal{K})^0$ if and only if $\langle \theta, \mathbf{z} \rangle \leq 0$ for every $\mathbf{z} \in \text{cone}(1 \oplus \mathcal{K})$. Observe that it suffices

to restrict to only the set generating the cone, that is $\theta \in \text{cone}(1 \oplus \mathcal{K})^0$ if and only if $\langle 1 \oplus \mathbf{x}', \mathbf{z} \rangle \leq 0$ for each $\mathbf{x}' \in \mathcal{K}$. Hence,

$$\begin{aligned} \ell(\mathbf{x}, \mathbf{f}) \in S &\iff \langle \langle \mathbf{f}, \mathbf{x} \rangle \oplus -\mathbf{f}, \mathbf{z} \rangle \leq 0 \quad \forall \mathbf{z} \in \text{cone}(1 \oplus \mathcal{K}) \\ &\iff \langle \langle \mathbf{f}, \mathbf{x} \rangle \oplus -\mathbf{f}, 1 \oplus \mathbf{x}' \rangle \leq 0 \quad \forall \mathbf{x}' \in \mathcal{K} \\ &\iff \langle \mathbf{f}, \mathbf{x} \rangle \leq \langle \mathbf{f}, \mathbf{x}' \rangle \quad \forall \mathbf{x}' \in \mathcal{K} \end{aligned}$$

Of course, this can be achieved by setting $\mathbf{x} = \arg \min_{\mathbf{x} \in \mathcal{K}} \langle \mathbf{f}, \mathbf{x} \rangle$, and hence we are done. \square

Proof of Proposition 2. First notice that we require the halfspace-satisfiability condition for S to ensure that the “halfspace oracle” in Algorithm 2 exists. Because we are selecting θ_t in $\mathcal{K} \subset \text{cone}(S)^0$, θ_t defines a halfspace containing S and hence we can use our halfspace oracle to find an \mathbf{x}_t satisfying $\langle \theta_t, \ell(\mathbf{x}_t, \mathbf{y}) \rangle \leq 0$ for every $\mathbf{y} \in \mathcal{Y}$.

To bound $D_T(\mathcal{A})$, which is the distance between the point $\frac{1}{T} \sum_{t=1}^T \ell(\mathbf{x}_t, \mathbf{y}_t)$ and the set S , we begin by instead bounding the distance to $\text{cone}(S)$. We can immediately apply Lemma 3 to obtain

$$\begin{aligned} \text{dist} \left(\frac{1}{T} \sum_{t=1}^T \ell(\mathbf{x}_t, \mathbf{y}_t), \text{cone}(S) \right) &= \max_{\theta \in \mathcal{K}} \left\langle \frac{1}{T} \sum_{t=1}^T \ell(\mathbf{x}_t, \mathbf{y}_t), \theta \right\rangle = \frac{1}{T} \max_{\theta \in \mathcal{K}} \left(- \sum_{t=1}^T \langle \mathbf{f}_t, \theta \rangle \right) \\ &\leq \frac{1}{T} \left(\sum_{t=1}^T \langle \mathbf{f}_t, \theta_t \rangle - \min_{\theta \in \mathcal{K}} \sum_{t=1}^T \langle \mathbf{f}_t, \theta \rangle \right) = \frac{1}{T} \text{Regret}_T(\mathcal{A}) \quad (10) \end{aligned}$$

where the first inequality follows by the halfspace oracle guarantee. Of course, if we let $S' \subset \mathbb{R}^{d-1}$ be the set S after removing the first coordinate, then we see by Lemma 4 that for any $\mathbf{z} \in \mathbb{R}^{d-1}$,

$$\text{dist}(1 \oplus \mathbf{z}, S) = \text{dist}(\mathbf{z}, S') \leq (1 + \|S'\|) \text{dist}(1 \oplus \mathbf{z}, \text{cone}(1 \oplus S')) \leq (1 + \|S\|) \text{dist}(1 \oplus \mathbf{z}, \text{cone}(S)). \quad (11)$$

By assumption, however, we can write $\frac{1}{T} \sum_{t=1}^T \ell(\mathbf{x}_t, \mathbf{y}_t) = 1 \oplus \mathbf{z}$ for some $\mathbf{z} \in \mathbb{R}^{d-1}$. Combining this with equations (10) and (11) finishes the proof. \square

Proof of Proposition 1. Applying Lemma 3 to the definition of $D_T(\mathcal{A})$ gives

$$D_T(\mathcal{A}) \equiv \text{dist} \left(\frac{1}{T} \sum_{t=1}^T \ell(\mathbf{x}_t, \mathbf{f}_t), S \right) = \max_{\mathbf{w} \in \text{cone}(1 \oplus \mathcal{K}), \|\mathbf{w}\| \leq 1} \left\langle \frac{1}{T} \sum_{t=1}^T \ell(\mathbf{x}_t, \mathbf{f}_t), \mathbf{w} \right\rangle \quad (12)$$

Notice that, in this optimization, we can assume w.l.o.g. that $\|\mathbf{w}\| = 1$, or $\mathbf{w} = \mathbf{0}$. In the former case we can write $\mathbf{w} = \frac{1 \oplus \mathbf{x}}{\|1 \oplus \mathbf{x}\|}$ for some $\mathbf{x} \in \mathcal{K}$, and we drop the latter case to obtain the inequality

$$\begin{aligned} D_T(\mathcal{A}) &\geq \max_{\mathbf{x} \in \mathcal{K}} \left\langle \frac{1}{T} \sum_{t=1}^T \ell(\mathbf{x}_t, \mathbf{f}_t), \frac{1 \oplus \mathbf{x}}{\|1 \oplus \mathbf{x}\|} \right\rangle = \frac{1}{T} \max_{\mathbf{x} \in \mathcal{K}} \frac{\left(\sum_{t=1}^T \langle \mathbf{f}_t, \mathbf{x}_t \rangle - \sum_{t=1}^T \langle \mathbf{f}_t, \mathbf{x} \rangle \right)}{\|1 \oplus \mathbf{x}\|} \\ &\geq \frac{\frac{1}{T} \left(\sum_{t=1}^T \langle \mathbf{f}_t, \mathbf{x}_t \rangle - \sum_{t=1}^T \langle \mathbf{f}_t, \mathbf{x}^* \rangle \right)}{\|1 \oplus \mathbf{x}^*\|} \geq \frac{\frac{1}{T} \text{Regret}_T(\mathcal{A})}{1 + \|\mathcal{K}\|}, \end{aligned}$$

where we set $\mathbf{x}^* := \arg \min_{\mathbf{x} \in \mathcal{K}} \sum_{t=1}^T \langle \mathbf{f}_t, \mathbf{x} \rangle$. \square

B A generalization of Blackwell to convex functions

Consider the following generalization of Blackwell to functions. In analogy to Blackwell, let:

1. A two-player game with functions as payoffs. For strategies i, j we have that the payoff of the game $l(i, j) \in S$ is a function (rather than a vector as in Blackwell).
2. S - set of functions $S : \{f : \mathbb{R}^d \mapsto \mathbb{R}\} \subseteq \mathcal{F}$.
3. "Halfspaces", which are characterized by $x \in \mathbb{R}^d$. The halfspace H_x contains all functions f such that $f(x) \leq 0$.
4. An oracle $O : H_x \mapsto p$ which maps a halfspace containing S , i.e. $\forall f \in S, f(x) \leq 0$, into a distribution over player strategies, such that the resulting loss function is contained inside the halfspace, i.e.

$$\forall j, l(O(x), j) = l(p, j) \in H_x$$

5. When talking of approachability we need a distance measure. If we think of $f(x)$ as the "inner-product" between f and x , then $K = \{x | f(x) \leq 0 \ \forall f \in S\}$, which is the "dual" set to S , is the set of all hyperplanes containing S . Our distance measure between a function f and S is then taken to be the maximal inner product with any hyper-plane containing S :

$$d(f, S) \equiv \max\{0, \max_{x \in K} f(x)\}$$

Note that this distance is zero for all members of S .

Then:

Theorem 7 (Blackwell generalization for functions). *Given an Oracle as above, the set S is approachable.*

The Blackwell method of proof is geometric in Nature, and it is not immediately clear how to generalize it to prove the above. However, using Online Convex Optimization, the proof is a simple generalization of the one in the previous sections:

Proof. Define the dual set to S as

$$K = \{x | f(x) \leq 0 \ \forall f \in S\}$$

Define the gain function for iteration t - f_t - as $f_t = l(p_t, j_t)$ where j_t is the adversary's strategy at iteration t , and p_t is given by the oracle as the mixed user strategy for the hyperplane parameterized by x_t .

Hence, iteratively the OCO algorithm generates an $x_t \in K$, which is then fed to the Oracle to obtain $p_t = O(x_t)$, which in turn defines the cost function for this iteration $f_t = l(p_t, j)$. The OCO low-regret theorem guaranties us that

$$\max_{x^*} \frac{1}{T} \sum_t f_t(x^*) - \frac{1}{T} \sum_t f_t(x_t) \leq \varepsilon_t \mapsto 0$$

By the guarantee provided by the oracle, we have that $f_t(x_t) \leq 0$, which combined with the above gives us:

$$d(\bar{f}, S) = \max_{x^*} \bar{f}(x^*) = \max_{x^*} \frac{1}{T} \sum_t f_t(x^*) \leq \varepsilon_t \mapsto 0$$

Which by our definition implies that the distance of the average gain function to the set S converges to zero. \square